

J. Clin. Chem. Clin. Biochem.

Vol. 25, 1987, pp. 901–905

© 1987 Walter de Gruyter & Co.
Berlin · New York

Report on a Joint European Quality Control Survey for Neonatal Thyrotropin Determinations (1986)

By G. Röhle, R. Kruse and W. J. Geilenkeuser

Institut für Klinische Biochemie der Universität Bonn

(Received July 2/September 25, 1987)

Summary: This paper presents the results of a joint European external quality control survey for thyrotropin determinations in blood dried on filter paper, carried out in 1986 in cooperation with several national quality control organizations. For the evaluation, 124 participants presented their individual diagnostic classifications in addition to their analytical results. Although, in relation to earlier studies of this kind, there was a significant improvement in interlaboratory precision the results still showed variance which depended on the analytical method and, possibly on the country in which it was performed. Regional differences were also evident in the diagnostic classifications.

Introduction

On the 5th. and 6th. December, 1984, a symposium was held in Lille, France on the subject of "Neonatal Screening and Quality Control". One of the results of the symposium was a plan to carry out annually a joint European quality control survey of screening analyses for hypothyroidism and phenylketonuria.

The first joint European external quality control survey of thyrotropin determinations was conducted by the French organization for external quality assessment in December 1985, the second by the "Deutsche Gesellschaft für Klinische Chemie" in June 1986. The UK External Quality Assessment Scheme for Thyroid-related Hormones is now preparing a third one for September 1987.

The present paper reports on the results of the second joint European quality control survey.

Materials and Methods

Control specimens

As samples, the participants each received 4 cards of filter paper (Schleicher & Schüll 2992) with 3 drops of dried blood each. The desired thyrotropin concentrations were attained by adding the 1st reference preparation 1974 (68/38) to whole blood containing low thyrotropin concentrations (haematocrit-corrected).

This resulted in the following concentrations, which were also taken as target values for the evaluation of the individual results:

Sample 1	13 mU/l
Sample 2	7 mU/l
Sample 3	25 mU/l
Sample 4	37 mU/l

Quality control survey

Four different samples each were sent to 153 laboratories in the following countries:

France	25	
Great Britain	30	
Italy	24	
Spain	16	(including one laboratory in Portugal)
Federal Republic of Germany	31	
Others	27	(Laboratories of 14 other countries that ordinarily participate directly in the collaborative surveys of the Deutsche Gesellschaft für Klinische Chemie)

The analyses were carried out from June 9 to 21, 1986.

Documentation of the results and findings

Together with control samples, the participants in the quality control surveys received a registration form. The results of two single determinations and their mean in mU per l of blood were

entered on these forms. In addition, a diagnostic classification of the analysis result was required:

- 1) normal
- 2) a case of congenital hypothyroidism possibly or
- 3) probably pathologic.

Further, when a commercially available test kit had been used for the determinations, the name of the kit producer was requested.

Evaluation

After the results had been collected, evaluation was then begun with the help of electronic data processing. The mean of the individual estimations for each double determination was data processed as the analysis result.

Changes made after the first evaluation of the survey had been sent out to the participants:

Four participants had cited the thyrotropin concentration in serum instead of in blood, as was indicated in the experiment instructions. As a result of information handed in at a later date, the readings from these participants were halved for this present evaluation. Moreover, now included are four additional participants whose results had arrived too late for initial evaluation.

In order to detect possible correlations, the results were divided into 2 sub-groups:

- 1) Those that had worked with the same commercial kits and
- 2) those from the same country (excepting a small number of laboratories from countries in the subgroup "others")

The median of the values served to establish the location parameter of the collective values, and the distance between the 25%–75% and the 16%–84% percentiles, respectively, served to describe the dispersion.

Results and Discussion

124 laboratories sent in their analysis results for evaluation. The analysis methods applied are based on the analytic principles for determining thyrotropin in dried blood that were developed in the seventies (1, 2). According to the particulars given by the participants, 21 of them used modifications they had developed themselves, 88 carried out the analysis with test kits from 12 different and named manufacturers, while 15 used other kits or gave no data (tab. 1).

In agreement with earlier studies (3, 4), there were again no essential differences between the target values and the medians of all the results (tab. 2). Inter-laboratory precision, however, had significantly improved, as can be seen by comparing the relative deviations of 16% and of 84% percentiles with the corresponding values from past external quality control surveys. This trend toward greater precision, already apparent as early as 1983, therefore continues for these semi-quantitative determinations (5).

An important reason for this improvement seems to be that the differences in results now depend less on which kit is used (6). If one takes for comparison the medians of results obtained but at least 3 participants with the same commercial kit, then for a specimen with 31 mU/l thyrotropin in 1982 (3) the greatest, relative negative deviation of a kit median from the

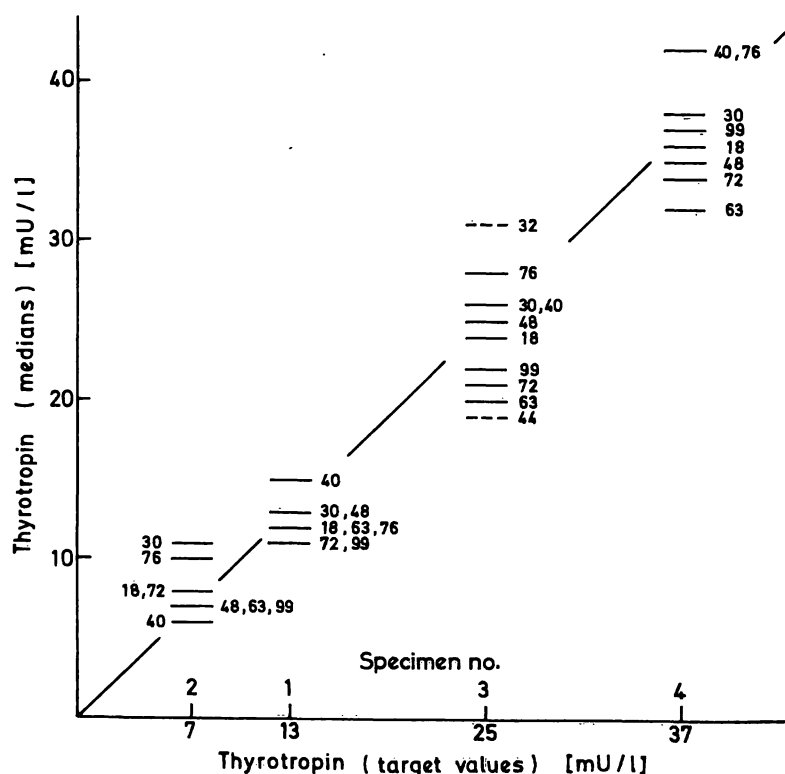


Fig. 1. Medians of results from at least six laboratories using the same commercial kits for 4 different thyrotropin specimens. The numbers in the figure are the code numbers according to table 1. For specimen 3 the medians of the kits No. 32 (3 laboratories) and No. 44 (4 laboratories) are given additionally (see text).

Tab. 1. Number of participating laboratories from different countries using different kits.

Commercial Kits	Code No.	France	F.R. Germany	Great Britain	Italy	Spain Portugal	Other countries
Becton-Dickinson	18		7	1	3		3
Behring	20		1				
CIS/CEA	30	18	1				
Mallinckrodt	32		1		2		
Corning	40		1	6		2	
Diagnostic Products	44				1	2	1
Henning	48		6				1
IRE	53						
LKB	63		6	3			7
Pharmacia	72			4			4
Roche	74					1	
Serono/Biodata	76		1		5		1
Other	98	3		4	1	5	2
Own methods	99		1	11	1	2	5

Tab. 2. Target values, medians (M), 16% and 84% percentiles as well as their relative deviations from the medians of the joint European thyrotropin survey 1986. For comparison the mean relative deviations calculated from 7 surveys in 1980 – 1982 are shown for the given concentrations (3).

Sample	Target value (mU/l)	Median (mU/l)	Percentiles (mU/l)		Relative deviations of percentiles			
			16%	84%	1986	1980–82 (mean)		
						M – 16%	84% – M	M – 16%
						M	M	M
1	13	12.1	8.74	16.6	–0.28	+ 0.37	–0.32	+ 0.50
2	7	7.99	5.11	12.0	–0.36	+ 0.50	–0.42	+ 0.79
3	25	23.8	18.2	29.5	–0.24	+ 0.24	–0.27	+ 0.39
4	37	36.5	29.0	44.0	–0.21	+ 0.21	–0.26	+ 0.39

target value was -0.32 and the greatest, relative positive deviation $+0.35$. In 1986, maximum relative deviations of -0.25 and $+0.25$ were found for the specimen with 25 mU/l, when – as is necessary for comparison with the former data – kits No. 32 (3 participants) and No. 44 (4 participants) were additionally taken into account (fig. 1).

However, a first glance still gives the impression that the type of kits used played a leading role in the results. Although the medians of the results from France agreed very well in 3 out of 4 cases with the target values, it is otherwise conspicuous that in France where 18 from 21 laboratories used kit No. 30 and in Italy where 5 from 13 laboratories preferred kit No. 76, the highest values, almost without exception, were recorded. As no comparative values for other known kits were available for France, it is not possible to decide whether higher values were due to the kit or to the region. The results from Italy – where different kits were used – at least do not exclude the possibility that regional factors, whatever

they may be, could have played a role. Of their 52 results, only 11 were below the medians of all the participants; and with kit No. 18 the medians for Italy were as much as 4 mU/l (specimen 3) higher than for the Federal Republic of Germany.

Much more homogeneous than the analysis results were the diagnostic classifications which the participants derived from them (fig. 2). For specimens 1, 2 and 4, only a few participants cited a classification – either “normal” or “possibly/probably pathological” – which deviated from the majority. These few divergent classifications could be due to grave, analytical errors.

It is a different case for the classification of the specimen in which the thyrotropin concentration (25 mU/l) is close to the decision limit, set most frequently at around 20 mU/l by the laboratories (3). Here, as a result of unavoidable analytic dispersion, there should have been a higher proportion of divergent classifications. It was therefore remarkable when a considerable number of French laboratories listed this

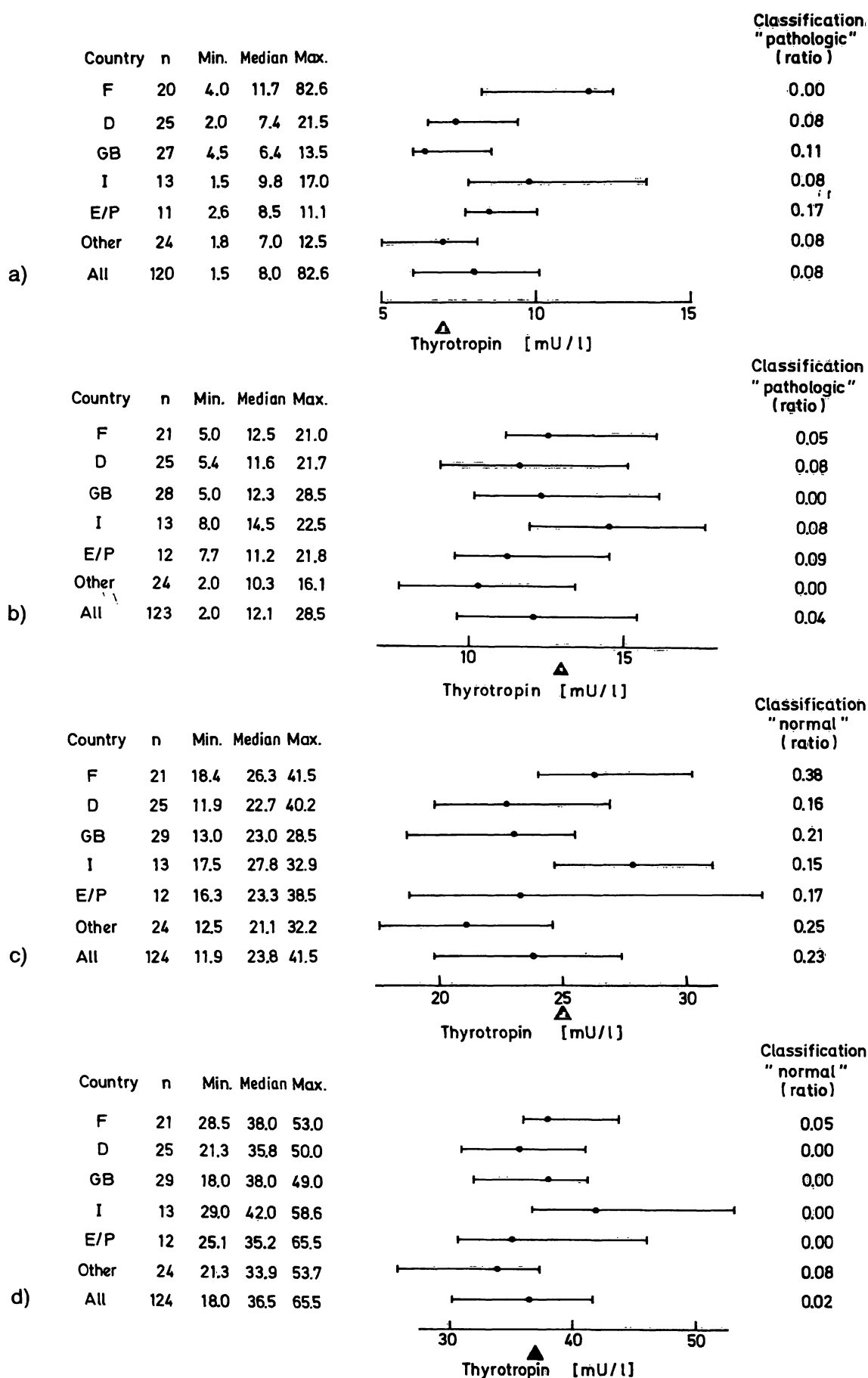


Fig. 2, Survey results of thyrotropin determinations established by laboratories from different countries plus the ratio of the diagnostic classifications that differ from the majority: "normal" and "possibly/probably pathologic" respectively.

Fig. 2a: specimen 2; fig. 2b: specimen 1; fig. 2c: specimen 3; fig. 2d: specimen 4.

(n = number; Min. = lowest value; Med. = median; Max. = highest value; —●— 25%, median and 75% percentile; ▲ location of target value. Countries: F = France; D = Federal Republic of Germany; GB = Great Britain; I = Italy; E/P = Spain/Portugal)

classification as "normal", although only one of them had found a value under 20 mU/l. The decision limit is apparently allowed a higher range in France than in most other countries; in the Federal Republic of Germany it is set at a maximum 20 mU/l. As there is no evidence that French screening often fails to

detect cases of congenital hypothyroidism, the question arises of whether a higher value should not be agreed on in the interest of coordinating the settings for decision limits. This was suggested in 1983 by *Wieland & v. Reuss*, in order to reduce the number of superfluous, data-seeking tests on patients.

References

1. Dussault, J. H., Parlow, A., Letarte, J., Guyda, H. & Laberge, C. (1976) *J. Pediat.* 89, 550–554.
2. Illig, R. & Rodriguez de Vera Roda, C. (1976) *Schweiz. Med. Wochenschr.* 106, 1676–1681.
3. Röhle, G., Voigt, U., Kruse, R. & Torresani, T. (1983) *J. Clin. Chem. Clin. Biochem.* 21, 813–821.
4. Röhle, G., Kruse, R. & Voigt, U. (1985) In: *Neugeborenen-Screening auf hereditäre Stoffwechselkrankheiten und endokrine Störungen* (Bickel, H. & Wachtel, U., eds.) Georg Thieme Verlag Stuttgart–New York, pp. 67–76.
5. Röhle, G. & Voigt, U. (1986) *Commun. Lab. Med.* 2, 81–85.
6. van Thiel, D., Marschner, I., Wood, W. G., Habermann, J. & Scriba, P. C. (1980) *J. Clin. Chem. Clin. Biochem.* 18, 807–816.
7. Wieland, O. H. & v. Reuss, K. (1983) *Dt. Ges. f. Klin. Chemie e. V. — Mitteilungen* 14, 95–96.

PD Dr. G. Röhle
Institut für Klinische Biochemie
Sigmund-Freud-Straße 25
D-5300 Bonn 1

